|  |
| --- |
| University of Queensland |
| Project Proposal |
| What is the minimal number of symbols required to estimate entropy accurately for dialogue and monologue |
|  |
| **Robert Cochran - 43135132** |
| **8/8/2018** |

|  |
| --- |
| [Type the abstract of the document here. The abstract is typically a short summary of the contents of the document. Type the abstract of the document here. The abstract is typically a short summary of the contents of the document.] |

Contents

[Research Questions 2](#_Toc521666087)

[What is the minimal symbol set required to produce accurate and distinct entropy calculations 2](#_Toc521666088)

[Creating minimal alphabets that can produce complex and extensive entropies 2](#_Toc521666089)

[allow for accurately ~~identify a speaker in a group through entropy?~~ 2](#_Toc521666090)

[~~Can we find high valued/information symbols that help to capture desired results more efficiently?~~ 2](#_Toc521666091)

[Introduction 2](#_Toc521666092)

[Addresses the significance of the research 2](#_Toc521666093)

[What have been the drivers of and motivations for this research? 2](#_Toc521666094)

[Why does this area need research or a research base in general? 2](#_Toc521666095)

[What then is the broad aim of the proposed research? 2](#_Toc521666096)

[Identification 2](#_Toc521666097)

[Previous Research / Literature Review 4](#_Toc521666098)

[Motivations 4](#_Toc521666099)

[Pauses as symbols in speech 4](#_Toc521666100)

[Classes of Pause symbols 4](#_Toc521666101)

[Goals: 5](#_Toc521666102)

[Entropy Calculation 5](#_Toc521666103)

[Conversational Analysis 6](#_Toc521666104)

[Dementia Analysis 7](#_Toc521666105)

[Theoretical Framework/Hypothesis 8](#_Toc521666106)

[Methods 8](#_Toc521666107)

[Project Plan 8](#_Toc521666108)

[To Do: 8](#_Toc521666109)

[Notes 8](#_Toc521666110)

[Bibliography 9](#_Toc521666111)

Thesis

# Research Questions

1. What is the minimal symbol set required to produce accurate and distinct entropy calculations
2. Creating minimal alphabets that can produce complex and extensive entropies
3. allow for accurately ~~identify a speaker in a group through entropy?~~
4. Can we find high valued/information symbols that help to capture desired results more efficiently?

<https://uq.edu.au/student-services/pdf/learning/research-proposal-sample-v2.pdf>

# Introduction

### Trouble and Repair

~~Conversational breakdowns in Dementia have been extensively researched to establish where exactly trouble starts occurring in conversations between People with Dementia (PWD) and their carers or loved ones. For PWD conversations serve not only as a polite, social courtesy with carers, but also act as a form of therapy as less communication leads to a deterioration in mental health to those suffering [needed citation]. Currently approaches to help repair conversations are done through user training, but a natural question to ask given the abundance of computational power available is if trouble detection can be automated, and if so, can it then automate or aid in repair as well?~~

~~Before repair can start, signs of trouble in speech must be studied to determine where exactly breakdowns start occurring in communication and conversation. Trouble Indicating Behaviours (TIB’s) in speech have been well studied to find key points in communication where these conversation breakdowns occur, both in normal conversation and conversations involving PWD. TIB’s are defined by Chenery et. al(1995) as conversational tools for a listener to “highlight points of trouble in understanding a message the speaker is intending”. Using conversational analysis, Chenery et al. have documented not only several TIB’s but also the repair techniques used to re-establish the conversation immediately after and their effectiveness [is this true].~~

~~Examples of TIB’s can be requesting for more or specific information, reprise/minimal dysfluency (“um”, “ah”, “yes”, “hmmm”, frequent or disruptive pauses). According to data taken from Chenery et al, PWD are more likely to rely on certain types of TIB’s such as minimal dysfluency or a lack of uptake in the conversation, both being characterized by frequent or disruptive pauses to the conversation. This make sense as Dementia causes trouble in forming or recollecting memories, making conversation difficult if they rely on certain types of information retrieval (i.e. short or long term).~~

~~Given that trouble can be characterized by pauses, it make sense to try to look for them as a sign of trouble??~~

~~What do we study?~~

~~How do we detect trouble indicating behaviours?~~

### ~~Automating Trouble Detection~~

~~Can we then AUTOMATE trouble detection?~~

~~A natural thought is considering if trouble detection can be automated, by doing so it can lead to automated repair techniques that allows technology to take some of the burden off the patients when they are having trouble, and carers when they need help. Tech can help by providing a list of possible keywords to the patient when they are having trouble recalling a word based upon what they are more likely to say in the context and their own personal vocabulary from previous conversations (solutions tailored to the patient).~~

~~…~~

~~…~~

~~Where do we start?~~

~~What are methods of detection? (neural networks (might be too complicated or computation heavy), entropy calculation (more flexible? and lightweight))~~

~~Start with simpliset case.~~

~~Communication breakdown has a severe impact on the lives of People with Dementia (PWD)[?], by understanding the works on trouble and repair it becomes possible to look at simple instances of speech and try to use them to help repair conversations that become lost or difficult for PWD. By analysing speech it’s possible to implement previous research done in trouble and repair and implement those findings to detect trouble automatically through speech analysis (natural language processing?).~~

~~If we study speech, what are we going to study to produce concrete results? Where do we start first? We need to start small and work our way up with as much confidence as possible.~~

~~Currently there are a lot of aspects of speech that can be studied as ‘trouble indicating behaviours’ that has potential for analysis. To study these effectively is a lot of work, but the potential is that once solid language markers are found in PWD, they can be used as areas of repair in conversations.~~

~~PWD need conversations, it is not only a form of therapy that is nice in the short term (having a conversation), but can help PWD feel fulfilled and able?~~

~~A lot of trouble and repair techniques are focused on administering~~

~~but this research … but could this analysis/detection/reporting be done through tech, can we detect trouble in speech?~~

~~Dementia has seen assistance from technology in several areas [https://www.dementia.org.au/resources/assistive-technology], but one of the most innovative and beneficial areas in which technology can assist is in speech and repairing communication breakdown using Conversational Analysis. Include Trouble and repair. Conversational Analysis. Include types of~~

~~Natural language processing (NLP) has several potential applications. Particularly in health where it can be important to measure the change in a speaker’s speech patterns to make note of potential declining health problems or to aid in communication break downs. Thus, an important feature of NLP is not only in its ability to analyse conversations in the present, but to also analyse them against previous conversations and measure change over a long-term period.~~

To make sure that it can accurately analyse conversations to the full breadth of meaning that is being conveyed, nlp needs to not only understand the semantic lexical content, but the implied, auxiliary content that can change or influence meaning taken at face value.

### ~~What have been the drivers of and motivations for this research?~~

~~Research into essential/minimal symbol sets for natural language processing is important because~~

1. ~~it can speed up calculations/estimations~~
2. ~~provide greater insight into the actual meaning of a conversation~~
3. ~~allows for greater nuance in applications built around NLP~~

### ~~Why does this area need research or a research base in general?~~

~~It’s important that research be done that isn’t influenced by biased direction in development or for exploitative outcomes. This is fairly new research that should be addressing the content as fully and correctly as possible now so that it’s not being built upon biased work. ??? Not sure how necessary this is, feels forced to be honest.~~

### ~~What then is the broad aim of the proposed research?~~

~~Dementia -> tracking change/identifying -> alerting -> entropy > symbol sets > symbol analysis~~

~~Being able to identify speakers through their vocabulary is important and can be an efficient an unobtrusive method of doing so. Being able to then detect variations or change in speech is also important as it allows for potential aid for dealing with patients who might be developing problems in their language processing.~~

This requires a system that can capture speech, symbolize it, find the entropy and return data quickly. Ultimately the symbolization/alphabet should be minimal such that entropy can be collected with a minimal number of samples needed to return data as quickly as possible. Determining that symbol set is the main focus of research for this project. **This symbol set (alphabet) would need to encapsulate the highest valued aspects of speech** ~~relative to the desired outcome of analysis for that purpose (i.e. a symbol set would differ for focussing on PWD versus someone with a speech impediment).~~ Understanding specifically what symbols deliver the most amount of information for analysis purposes is crucial.

~~In this project speech patterns will be used to identify, in the future this could be used as a means for drawing meaning from a speech to provide a greater range and depth of analysis.~~

# Previous Research / Literature Review

## Using Speaking Pattern as Identification

## Conversational Analysis

Conversations carry with them a remarkably rich amount of information in both the direct lexical semantics delivered by speakers, but also their use of auxiliary forms of communication (e.g. shifts in tone, pitch, pause lengths, etc…) that provide inherent meaning and can modify the message being delivered.

~~To accurately analyse speech there must be specific, meaningful markers that are common enough in speech that these markers are reliable to find but also carry with them enough information that they provide some insightful or meaningful piece of information about either the speaker themselves or what the speaker is saying.~~

~~To make entropy estimation/calculation as efficient as possible, the right symbols need to be chosen. This means rethinking the set of symbols that are being used to analyse speech and try to find new symbols to look for that carry more information than through word analysis alone/that isn’t being implicitly said through a transcript.~~

~~These auxiliary forms of communication can be inflection, intonation, deliberate pausing, etc. that all carry inherent meaning with them. Because of this, they are used by speakers in very specific ways that can be thought of as being part of their vocabulary. This serves to identify certain speakers from others.~~

~~This can range from looking at the pauses that someone makes in their speech to their pitch, utterance lengths and to tonal shifts. All of these are areas in which we communicate and deliver meaning but are not necessarily well studied for their meaning. Additionally, they can be used to identify certain speakers (as we rely on them as part of our communicational behaviour like we would certain phrases or choice of words, we stick to the ones we like).~~

~~Semiotics is the study of symbols and meaning, this could be useful in determining further areas of symbols to analyse.~~

~~In this proposal accurate and distinct are being used to mean the possible entropy calculations should make it possible to fulfil and be useful in two different use cases. In this case, accurate is describing an entropy set that has a small amount of variance for each person. Distinct means each person should also not overlap significantly at all for entropy sets. So a total set of entropy values should allow for accurate distinction between speakers. This means the symbol set should support a minimum number N amounts of users to be accurately identified.~~

## ~~Pauses as symbols in speech~~

~~Pauses in speech can deliver a few intended or unintended meanings with them. For a speaker to respond in .5 seconds shows little care in that they are not taking the time to think and process the other speaker’s comments, too quick a response shows no thoughtful contemplation over what was said. Too long and it can show the speaker is not properly paying attention and has become disinterested.~~

~~Because pauses carry meaning it changes the way they are used, and thus could potentially serve as a marker for a conversational tool for a speaker to rely on, thus becoming part of their vocabulary.~~

~~It takes a minimum amount of time for a person to listen, take on board and respond accordingly. I think it takes 700ms just to comprehend the sentence that was said?? Given that there is a biological limit, it makes sense why pauses that are too short can come across as rude. But a pause can also occur when a conversation topic has ended and a new topic is being suggested, which will not carry the same time parameters to illicit meaning that the previous pause was able to.~~

~~But parameters that make sense for one region won’t make sense for another. Pause meanings can change with each culture and each person, the Japanese have “one of the shortest conversational replies … often answering before the conversational turn is over”[ Econimist, pauses]. This is not meant as a rude gesture, on the contrary as it helps move on the conversation along. Whereas in Finland it is customary to finish sentences with length pauses. With such wide culture~~

## ~~Classes of Pause symbols~~

~~Pauses in speech can be things like inner pauses (where a speaker has a brief pause to collect thoughts or let another person speak???). They can also be the time between a speaker stopping and another speaker picking up the conversation. This form of pause can carry with it a range of potential lengths, varying from 500ms to 5 seconds. This is a form of communication that can change the message being delivered or say something about the speaker’s intention or feelings or engagement in the conversation that isn’t being explicitly said.~~

~~We want to symbolize pauses, that means coming up with classes of pauses for the different areas in speech where they occur, and boundaries between which we can quantify them with (e.g. .5 second to 5 seconds). This will help find pauses that are the most likely or meaningful to the conversation and pauses that carry little meaning or information with them. The latter are pauses that can essentially be lumped together as they~~

~~So a pause can be after an overtake from speaker B and the return to speaker A, there might be a pause in-between B and A, that pause length can carry various meanings. It could be an indication of someone not paying attention to what was said (very quick return to speaking, i.e. short pause length), it could be a polite pause (.8-9), it could also mean the person is thinking about what was said and taking in the meaning of it to influence what they will say next (longer pause length). This pause class could then carry 3 symbols (possibly).~~

~~Google (Response time conversations)~~

[~~http://www.speech.kth.se/prod/publications/files/3859.pdf~~](http://www.speech.kth.se/prod/publications/files/3859.pdf)

[~~https://www.theatlantic.com/science/archive/2016/01/the-incredible-thing-we-do-during-conversations/422439/~~](https://www.theatlantic.com/science/archive/2016/01/the-incredible-thing-we-do-during-conversations/422439/)

[~~http://theconversation.com/awkward-pauses-in-online-calls-make-us-see-people-differently-26073~~](http://theconversation.com/awkward-pauses-in-online-calls-make-us-see-people-differently-26073)

[~~https://www.sciencedirect.com/science/article/pii/S0167639311001580~~](https://www.sciencedirect.com/science/article/pii/S0167639311001580)

[~~https://www.theatlantic.com/science/archive/2016/01/the-incredible-thing-we-do-during-conversations/422439/~~](https://www.theatlantic.com/science/archive/2016/01/the-incredible-thing-we-do-during-conversations/422439/)

[~~https://www.sciencedirect.com/science/article/pii/S1364661315002764~~](https://www.sciencedirect.com/science/article/pii/S1364661315002764)

[~~https://ac.els-cdn.com/S0095447010000628/1-s2.0-S0095447010000628-main.pdf?\_tid=c8148e1b-ddad-4986-9c4f-04fdd454257d&acdnat=1533785742\_707951e6b85bedda11decfbb6ea03f68~~](https://ac.els-cdn.com/S0095447010000628/1-s2.0-S0095447010000628-main.pdf?_tid=c8148e1b-ddad-4986-9c4f-04fdd454257d&acdnat=1533785742_707951e6b85bedda11decfbb6ea03f68)

[~~https://www.economist.com/books-and-arts/2017/12/14/the-importance-of-pauses-in-conversation~~](https://www.economist.com/books-and-arts/2017/12/14/the-importance-of-pauses-in-conversation)

~~How We Talk – NickEnfield~~

# Methods

## Benchmark Criteria

~~Creating a symbol set is vague and has a large potential to meet a variety of criteria, it needs to be able to achieve some purpose or have some kind of benchmark to measure up against. To test that the symbol set has achieved its goal, two sets will be carried out to determine the effectiveness and usefulness of such an alphabet.~~

**~~Case A:~~** ~~To use Shannon entropy as a marker to rank speakers in a group for identification and analysis purposes.~~

**~~Case B:~~** ~~To determine whether a given person’s current speech is significantly departing from that speaker’s expected norm.~~

## Goals:

**Overarching place in the field:** **Florence Project** - Aiding communication through technology. Analysing, tracking of recurrent patterns in conversation of PWD.

**Overarching project: Andrew’s work** - Detecting change in conversation or detecting when a meaningful pause has been made that shows conversation ending that is context specific (i.e. based on the person speaking), trying to aid in establishing early warning signs of early onset dementia or aiding in communicating with

**My project goals:** Use symbols to detect changes in speech in relation to pausing through text currently? What about speech? Or possibly a transcript? We want speech to analyse but is that too difficult, should we use a transcript first? Implement Andrews papers.

Related keywords: Dementia, Conversation, Recurrent conversational problems, Analysis and identification of this through text (Potentially transcript), Communication breakdown, symbols?, backchanneling, communication pauses, entropy in language

## Entropy Calculation

**Determining the Number of Samples Requires to Estimate Entropy in Natural Sequences** [[1](#Bac18)]

**Source:** <https://arxiv.org/pdf/1805.08929.pdf>

**Summary:** Given only a short sample size to compute entropy on a natural language, this provides a benchmark for reliable entropy calculation.

**Relevance:** This provides benchmarks for the minimum number of symbols the data should have to make correct conclusions about the entropy of those samples when analysing it.

**Fast Entropy Estimation for Natural Sequences** [[2](#Bac181)]

**Source:** <https://arxiv.org/abs/1805.06630>

**Summary:** Method for estimating the Entropy of a given message requiring significantly less samples than the alphabet size (e.g. not relying on slow histogram approaches) with minimal error.

**Relevance:** In trying to determine a speaker from a group (and whether they are departing from the norm), it’s important to be able to determine the entropy for given messages when only small amounts of samples are present. It’s this entropy calculation that allows for analysis and quantification of a person’s speech.

**Simple entropy estimator for small datasets** [[3](#Mon12)]

**Summary:** Given a small data set, it is not necessary to use a histogram approach to estimate entropy; it can be estimated through “character coincidences” needing only a small number of samples (relative to the alphabet size).

**Relevance:** This is important for developing a quick estimate of entropy calculation for a given sample set and provides the mathematical foundation outlined in the Fast Entropy paper above.

**Shannon Entropy Paper**

**Summary:** Defines rules for analysing and measuring entropy of language using statistics.

**Relevance:** Sets the foundation of research for all Information Theory texts.

## Conversational Analysis

**PauseCode: Computational Conversation Timing Analysis**

**Summary:** Looks at the important role pauses play in speech and the information they provide through their use.

**Relevance:** When determining what symbols to look for in speech, this paper outlines the impact a pause can have on illuminating a speaker’s interest, engagement or potential trouble in the conversation. Pause lengths that are too long can show the speaker is not engaged or having trouble following through the conversation (e.g. spending significant time remembering words). By using software to document and highlight meaningful pauses, tracking signs of trouble can become an automatic process.

**Conceptual Recurrence Plots: Revealing Patterns in Human Discourse** [[4](#Ang12)]

**Source:** <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5887327&isnumber=6180049&tag=1>

**Summary:** Visualises conversations by colour coding recurrent patterns on a semantic/conceptual level rather than a term based level (i.e. the same idea versus the same word). Salton is used to build the underlying semantic model.

**Relevance:** This helps with identifying when certain topics are being rediscussed continuously because of a breakdown in conversational ability for PWD. Reliance on a lot of back-channelling can be seen, as well as how deep a conversation is versus when not a lot of progress is being made to establish meaningful conversation (e.g. changing topics). Meaningful conversation is important for PWD and using visualisation techniques to track what does and doesn’t work is important. Visualisation can be quantified by the shading, frequency and the number of places that recurrences take place in the conversation.

**Human Communication, Quantifying Multi-participant recurrence (MPR) metrics** <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6161608>

Summary:

Relevance:

## Dementia Analysis

**Visualising dementia conversations** – Daniel, Janet, Helen: <https://www.cambridge.org/core/services/aop-cambridge-core/content/view/94641FE0FDBC9F84F38C28B5480F2E13/S0144686X13000640a.pdf/visualising_conversations_between_care_home_staff_and_residents_with_dementia.pdf>

**Summary:** Using Discursis to analyse the conversational patterns that emerge between PWD's and their carers.

**Relevance:** This provides documentation of further symbols to look for when estimating entropy and trying to pinpoint the presence of dementia through text or speech, and ways of helping those conversations through specific prompts.

**Automated Examination of Dementia Conversation** – Daniel, Janet, Helen: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0144327>

Summary:

Relevance:

**Trouble and Repair:** Helen Chenery, Janet Wiles

**Source:** <https://www.tandfonline.com/doi/pdf/10.1080/026870399402181>

**Summary:** Analysis of common Trouble Indicating Behaviours (TIB's) of patients with Senile Dementia of the Alzheimer's type (SDAT) and the effectiveness of using known repair types to aid in communication with their carer’s.

**Relevance:** Identifying a structured list of typical breakdowns for People with Dementia (PWD) in communication. This aims to aid in potential new paths of treatment or aid by identifying what patterns of speech work with PWD and what doesn’t. This can also be used to help with identifying valuable symbols or patterns in speech to look for.

PWD/SDAT (Senile dementia of the Alzheimer's type) conversational traits:

1. Shorter conversational turns,
2. called or regular prompts from the interviewer,
3. increased amount of reference errors,
4. missing elements in conversational turns of SDAT subjects,
5. more topic initiation and unexpected topic shifts (due to failure to continue and repetition of an idea.
6. Incoherent, meaningless or vague
7. Trouble Indicating Behaviours:
   1. Lack of uptake/Lack of continuation
   2. Reprise/Minimal dysfluency
   3. Pauses
   4. Request for repetition
   5. Not requesting specific information
   6. No hypothesis formation

**An analysis of trouble and repair in the natural conversations of people with dementia of the Alzheimer's type**

<https://www-tandfonline-com.ezproxy.library.uq.edu.au/doi/pdf/10.1080/026870399402181?needAccess=true>

**Signals and Systems: MIT Lectures**

<https://www.youtube.com/watch?v=-FHm2pQmiSM>

**Summary:** Looking at complex systems in abstract terms to simplify to just input and output  
**Relevance:** is this just background research or can I reference it as lit review

It abstracts away unnecessary elements of a problem. Looks at signals as input/output and the relationships between them in terms of a function for studying phenomena. Signal is an x-dimensional function (of time or possibly not, just need a function to model our movement with independent and dependent variables), can be multidimensional.

For our research we are looking at something that is both continuous time (CT) signal and a discrete time (DT) signal, CT in the actual verbal input from the speaker, but will be DT in the computational aspect of recording and analysing. How to convert from CT to DT and what do we need to capture and what is the cost for storing everything? What’s the slowest sample rate we can have?

Functions behave as they always will, they’re bound to their output from their input, but signals and systems is looking at manipulating how their information is sent back to us in a new order essentially. So if it’s f(2x) we are getting every second f(x) point, thereby squishing the data down to get half as much. F(-x) will give us a flipped version of what we are looking at, sending the information backwards, reverse order. F(x-250) is making the timeline wait longer to get to the 250 point, you are pushing it back to happen later (further into the x axis). This is essentially looking at how to manipulate functions given only the input and the output to change, how can we modify the data we’re getting back.

Are we actually just trying to see how to approximate a new function from a given function?

**Discursis**: Discursis would be used to potentially map out the pauses in communication

<http://discursis.com/index.php/news-and-research/>

**CalPy** - Software used to aid in this <https://github.com/YvonneYYu/calpy>

# Theoretical Framework/Hypothesis

# Methods

## Long term

Analyse and class pauses. Then move onto tonal shifts, pitch shifts, etc… and make a entropy vector.

## Short Term

Run experiments through CalPy to find pause lengths in speech.

To determine meaningful pauses

Outline: Send Audio to Calpy, make plots and determine Symbols, use those symbols to get Entropy

Run simple estimates in matlab, like entropy calculations

# Notes

3-4 weeks have a research

1. Since entropy is used more as a means of calculating a minimum compression, could there be problems with using it as an index? Could we incorrectly assess something as having less entropy than we think because of bad or few samples?
   1. Think Andrew's work covers this with the min number of samples needed

# Bibliography

x

|  |  |
| --- | --- |
| [1] | Andrew D Back, Daniel Angus, and Janet Wiles. (2018, May) Determining the Number of Samples Required to Estimate Entropy in Natural Sequences. |
| [2] | Andrew D Back, Daniel Angus, and Janet Wiles. (2018, May) Fast Entropy Estimation for Natural Sequences. |

x